Service Order Accuracy	$50	

: SEEM Submetrics

1.

Tier 1 Submetrics

Table contains a list of Tier 1 submetrics.

Table Tier 1 Submetrics (Continued)

Item No.	Submetric
1. 1	Loop Makeup - Response Time - Manual
2. 2	Loop Makeup - Response Time - Electronic
3.	Acknowledgement Message Timeliness
4.	Acknowledgement Message Completeness
5.	Percent Flow-Through Service Requests (Detail)
6.	Reject Interval
7.	Firm Order Confirmation Timeliness
8.	Firm Order Confirmation and Reject Response Completeness - Fully Mechanized
9.	Percent Missed Installation Appointments - Resale POTS
10.	Percent Missed Installation Appointments - Resale Design
11.	Percent Missed Installation Appointments - UNE Loop and Port Combinations
12.	Percent Missed Installation Appointments - UNE Loops
13.	Percent Missed Installation Appointments - UNE xDSL
14.	Percent Missed Installation Appointments - UNE Line Sharing
15.	Percent Missed Installation Appointments - Local IC Trunks
16.	Average Completion Interval - Resale POTS
17.	Average Completion Interval - Resale Design
18.	Average Completion Interval - UNE Loop and Port Combinations
19.	Average Completion Interval - UNE Loops
20.	Average Completion Interval - UNE xDSL
21.	Average Completion Interval - UNE Line Sharing
22.	Average Completion Interval - Local IC Trunks
23.	Coordinated Customer Conversions Interval - Unbundled Loops
24.	Coordinated Customer Conversions - Hot Cut Timeliness Percent within interval - UNE Loops
25.	Coordinated Customer Conversions - Percent Provisioning Troubles Received within 7 days of a completed service order - UNE Loops
26.	Cooperative Acceptance Testing - Percent of xDSL Loops Tested
27.	Percent Provisioning Troubles within 30 days of Service Order Completion - Resale POTS
28.	Percent Provisioning Troubles within 30 days of Service Order Completion - Resale Design
29.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE Loop and Port Combinations
30.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE Loops
31.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE xDSL
32.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE Line Sharing
33.	Percent Provisioning Troubles within 30 days of Service Order Completion - Local IC Trunks
34.	LNP - Percent Missed Installation Appointments - LNP
35.	LNP - Average Disconnect Timeliness Interval - LNP
36.	Missed Repair Appointments - Resale POTS
37.	Missed Repair Appointments - Resale Design

38.	Missed Repair Appointments - UNE Loop and Port Combinations	
39.	Missed Repair Appointments - UNE Loops	
40.	Missed Repair Appointments - UNE xDSL	
41.	Missed Repair Appointments - UNE Line Sharing	
42.	Missed Repair Appointments - Local IC Trunks	
43.	Customer Trouble Report Rate - Resale POTS	
44.	Customer Trouble Report Rate - Resale Design	
45.	Customer Trouble Report Rate - UNE Loop and Port Combinations	
46.	Customer Trouble Report Rate - UNE Loops	
47.	Customer Trouble Report Rate - UNE xDSL	
48.	Customer Trouble Report Rate - UNE Line Sharing	
49.	Customer Trouble Report Rate - Local IC Trunks	
50.	Maintenance Average Duration - Resale POTS	
51.	Maintenance Average Duration - Resale Design	
52.	Maintenance Average Duration - UNE Loop and Port Combinations	
53.	Maintenance Average Duration - UNE Loops	
54.	Maintenance Average Duration - UNE xDSL	
55.	Maintenance Average Duration - UNE Line Sharing	
56.	Maintenance Average Duration - Local IC Trunks	
57.	Percent Repeat Troubles within 30 days - Resale POTS	
58.	Percent Repeat Troubles within 30 days - Resale Design	
59.	Percent Repeat Troubles within 30 days - UNE Loop and Port Combinations	
60.	Percent Repeat Troubles within 30 days - UNE Loops	
61.	Percent Repeat Troubles within 30 days - UNE xDSL	
62.	Percent Repeat Troubles within 30 days - UNE Line Sharing	
63.	Percent Repeat Troubles within 30 days - Local IC Trunks	
64.	Invoice Accuracy	
65.	Mean Time to Deliver Invoices	
66.	Usage Data Delivery Accuracy	
67.	Trunk Group Performance - CLEC Specific	
68.	Collocation Percent of Due Dates Missed	

1.

Tier 2 Submetrics

Table contains a list of Tier 2 submetrics.

Table Tier 2 Submetrics (Continued)

Item No.	Tier 2 Sub Metrics
1. 1	Average Response Time - Pre-Ordering/Ordering
2. 2	Interface Availability - Pre-Ordering/Ordering
3.	Interface Availability - Maintenance & Repair
4.	Loop Makeup - Response Time - Manual
5.	Loop Makeup - Response Time - Electronic
6.	Acknowledgement Message Timeliness - EDI
7.	Acknowledgement Message Timeliness - TAG
8.	Acknowledgement Message Completeness EDI
9.	Acknowledgement Message Completeness TAG
10.	Percent Flow-through Service Requests (Summary)
11.	Reject Interval
12.	Firm Order Confirmation Timeliness
13.	Firm Order Confirmation and Reject Response Completeness - Fully Mechanized
14.	Percent Missed Installation Appointments - Resale POTS
15.	Percent Missed Installation Appointments - Resale Design
16.	Percent Missed Installation Appointments - UNE Loop and Port Combinations
17.	Percent Missed Installation Appointments - UNE Loops
18.	Percent Missed Installation Appointments - UNE xDSL
19.	Percent Missed Installation Appointments - UNE Line Sharing
20.	Percent Missed Installation Appointments - Local IC Trunks
21.	Average Completion Interval - Resale POTS
22.	Average Completion Interval - Resale Design
23.	Average Completion Interval - UNE Loop and Port Combinations
24.	Average Completion Interval - UNE Loops
25.	Average Completion Interval - UNE xDSL
26.	Average Completion Interval - UNE Line Sharing
27.	Average Completion Interval - Local IC Trunks
28.	Coordinated Customer Conversions Interval - Unbundled Loops
29.	Coordinated Customer Conversions - Hot Cut Timeliness Percent within interval - UNE Loops
30.	Coordinated Customer Conversions - Percent Provisioning Troubles Received within 7 days of a completed service order - UNE Loops
31.	Cooperative Acceptance Testing - Percent xDSL Loops Tested
32.	Percent Provisioning Troubles within 30 days of Service Order Completion - Resale POTS
33.	Percent Provisioning Troubles within 30 days of Service Order Completion - Resale Design
34.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE Loop and Port Combinations
35.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE Loops
36.	Percent Provisioning Troubles within 30 days of Service Order Completion - UNE xDSL
37.	Provisioning Troubles within 30 days of Service Order Completion - UNE Line Sharing
38.	Percent Provisioning Troubles within 30 days of Service Order Completion - Local IC Trunks
39.	LNP - Percent Missed Installation Appointments
40.	LNP - Average Disconnect Timeliness Interval
41.	Missed Repair Appointments - Resale POTS

42.	Missed Repair Appointments - Resale Design	
43.	Missed Repair Appointments - UNE Loop and Port Combinations	
44.	Missed Repair Appointments - UNE Loops	
45.	Missed Repair Appointments - UNE xDSL	
46.	Missed Repair Appointments - UNE Line Sharing	
47.	Missed Repair Appointments - Local IC Trunks	
48.	Customer Trouble Report Rate - Resale POTS	
49.	Customer Trouble Report Rate - Resale Design	
50.	Customer Trouble Report Rate - UNE Loop and Port Combinations	
51.	Customer Trouble Report Rate - UNE Loops	
52.	Customer Trouble Report Rate - UNE xDSL	
53.	Customer Trouble Report Rate - UNE Line Sharing	
54.	Customer Trouble Report Rate - Local IC Trunks	
55.	Maintenance Average Duration - Resale POTS	
56.	Maintenance Average Duration - Resale Design	
57.	Maintenance Average Duration - UNE Loop and Port Combinations	
58.	Maintenance Average Duration - UNE Loops	
59.	Maintenance Average Duration - UNE xDSL	
60.	Maintenance Average Duration - UNE Line Sharing	
61.	Maintenance Average Duration - Local IC Trunks	
62.	Percent Repeat Troubles within 30 days - Resale POTS	
63.	Percent Repeat Troubles within 30 days - Resale Design	
64.	Percent Repeat Troubles within 30 days - UNE Loop and Port Combinations	
65.	Percent Repeat Troubles within 30 days - UNE Loops	
66.	Percent Repeat Troubles within 30 days - UNE xDSL	
67.	Percent Repeat Troubles within 30 days - UNE Line Sharing	
68.	Percent Repeat Troubles within 30 days - Local IC Trunks	
69.	Invoice Accuracy	
70.	Mean Time to Deliver Invoices	
71.	Usage Data Delivery Accuracy	
72.	Trunk Group Performance - Aggregate	
73.	Collocation Percent of Due Dates Missed	
74.	Timeliness of Change Management Notices	
75.	Timeliness of Documents Associated with Change	
76.	Service Order Accuracy - Resale Residence	
77.	Service Order Accuracy - Resale Business	
78.	Service Order Accuracy - Resale Design	
79.	Service Order Accuracy - UNE Specials (Design)	
80.	Service Order Accuracy UNE (Non-design)	

1

Tier 3 Submetrics

Table contains a list of Tier 3 submetrics.

Table Tier 3 Submetrics (Continued)

Item No.	Tier 3 Sub Metrics
1 1.	Percent Missed Installation Appointments - Resale POTS
2 2.	Percent Missed Installation Appointments - Resale Design
3	Percent Missed Installation Appointments - UNE Loop
4	Percent Missed Installation Appointments - UNE Loop & Port Combo
5	Percent Missed Installation Appointments - UNE xDSL (ADSL, HDSL, UCL)
6	Percent Missed Installation Appointments - UNE Line Sharing
7	Percent Missed Installation Appointments - Interconnection Trunks
8	Average Completion Interval (OCI) & Order Completion Interval Distribution - Resale POTS
9	Average Completion Interval (OCI) & Order Completion Interval Distribution - Resale Design
10	Average Completion Interval (OCI) & Order Completion Interval Distribution - UNE Loop & Port Combo
11	Average Completion Interval (OCI) & Order Completion Interval Distribution - UNE xDSL (ADSL, HDSL, UCL)
12	Average Completion Interval (OCI) & Order Completion Interval Distribution - UNE Line Sharing
13	Average Completion Interval (OCI) & Order Completion Interval Distribution - Interconnection Trunks
14	Missed Repair Appointments - Resale POTS
15	Missed Repair Appointments - Resale Design
16	Missed Repair Appointments - UNE Loop + Port Combo
17	Missed Repair Appointments - UNE Loops
18	Missed Repair Appointments - UNE xDSL
19	Missed Repair Appointments - UNE Line Sharing
20	Missed Repair Appointments - Interconnection Trunks
21	Invoice Accuracy
22	Mean Time To Deliver Invoices
23	Trunk Group Performance - Aggregate
24	Collocation Percent of Due Dates Missed
25	Timeliness of Change Management Notices
26	Timeliness of Documents Associated with Change

: Statistical Properties and Definitions

Statistical Methods for BellSouth Performance Measure Analysis

1. ~~1.~~Necessary Properties for a Test Methodology

The statistical process for testing if competing local exchange carriers (CLECs) customers are being treat equally with BellSouth (BST) customers involves more than just a mathematical formula. Three key elements need to be considered before an appropriate decision process can be developed. These are

- the type of data,
- the type of comparison, and
- the type of performance measure.

Once these elements are determined a test methodology should be developed that complies with the following properties.

- ~~Like-to-Like Comparisons.~~*Like-to-Like Comparisons* – When possible, data should be compared at appropriate levels, e.g. wire center, time of month, dispatched, and residential, new orders. The testing process should:
  - Identify variables that may affect the performance measure.
  - Record these important confounding covariates.
  - Adjust for the observed covariates in order to remove potential biases and to make the CLEC and the ILEC units as comparable as ~~possible.~~
  - possible.
- ~~Aggregate Level Test Statistic.~~*Aggregate Level Test Statistic* – Each performance measure of interest should be summarized by one overall test statistic giving the decision maker a rule that determines whether a statistically significant difference exists. The test statistic should have the following properties.
  - The method should provide a single overall index, on a standard scale.
  - If entries in comparison cells are exactly proportional over a covariate, the aggregated index should be very nearly the same as if comparisons on the covariate had not been done.
  - The contribution of each comparison cell should depend on the number of observations in the cell.
  - Cancellation between comparison cells should be limited.
  - The index should be a continuous function of the ~~observations.~~
  - observations.

- ~~Production Mode Process.~~ *Production Mode Process* – The decision system must be developed so that it does not require intermediate manual intervention, i.e. the process must be a "black box."
  - Calculations are well defined for possible eventualities.
  - The decision process is an algorithm that needs no manual intervention.
  - Results should be arrived at in a timely manner.
  - The system must recognize that resources are needed for other performance measure-related processes that also must be run in a timely manner.
  - The system should be auditable, and adjustable over ~~time.~~ time.

- ~~Balancing.~~ *Balancing* – The testing methodology should balance Type I and Type II Error probabilities.
  - P(Type I Error) = P(Type II Error) for well defined null and alternative hypotheses.
  - The formula for a test's balancing critical value should be simple enough to calculate using standard mathematical functions, i.e. one should avoid methods that require computationally intensive techniques.
  - Little to no information beyond the null hypothesis, the alternative hypothesis, and the number of observations should be required for calculating the balancing critical value.

~~**Trimming. Trimming of**~~ *Trimming* – Removing extreme observations from BellSouth and CLEC distributions is needed in order to ensure that a fair comparison is made between performance measures. Three conditions are needed to accomplish this goal. These are:

- Trimming should be based on a general rule that can be used in a production setting.
- Trimmed observations should not simply be discarded; they need to be examined and possibly used in the final decision making process.
- Trimming should only be used on performance measures that are sensitive to "outliers."

Measurement Types

The performance measures that will undergo testing are of four types:

- means
- proportions,
- rates, and
- ratio

While all four have similar characteristics, proportions and rates are derived from count data while means and ratios are derived from interval measurements.

1. ~~II.~~ Testing Methodology – The Truncated Z

Many covariates are chosen in order to provide deep comparison levels. In each comparison cell, a Z statistic is calculated. The form of the Z statistic may vary depending on the performance measure, but it should be distributed approximately as a standard normal, with mean zero and variance equal to one. Assuming that the test statistic is derived so that it is negative when the performance for the CLEC is worse than for the ILEC, a positive truncation is done – i.e. if the result is negative it is left alone, if the result is positive it is changed to zero. A weighted average of the truncated statistics is calculated where a cell weight depends on the volume of BST and CLEC orders in the cell. The weighted average is re-centered by the theoretical mean of a truncated distribution, and this is divided by the standard error of the weighted average. The standard error is computed assuming a fixed effects model.

Proportion Measures

For performance measures that are calculated as a proportion, in each adjustment cell, the truncated Z and the moments for the truncated Z can be calculated in a direct manner. In adjustment cells where proportions are not close to zero or one, and where the sample sizes are reasonably large, a normal approximation can be used. In this case, the moments for the truncated Z come directly from properties of the standard normal distribution. If the normal approximation is not appropriate, then the Z statistic is calculated from the hypergeometric distribution. In this case, the moments of the truncated Z are calculated exactly using the hypergeometric probabilities.

Rate Measures

The truncated Z methodology for rate measures has the same general structure for calculating the Z in each cell as proportion measures. For a rate measure, there are a fixed number of circuits or units for the CLEC, n_{2j} and a fixed number of units for BST, n_{1j}. Suppose that the performance measure is a "trouble rate." The modeling assumption is that the occurrence of a trouble is independent between units and the number of troubles in n circuits follows a Poisson distribution with mean ~~λn~~ λ_n where ~~λ~~ λ is the probability of a trouble in 1 circuit and n is the number of circuits.

In an adjustment cell, if the number of CLEC troubles is greater than 15 and the number of BST troubles is greater than 15, then the Z test is calculated using the normal approximation to the Poisson. In this case, the moments of the truncated Z come directly from properties of the standard normal distribution. Otherwise, if there are very few troubles, the number of CLEC troubles can be modeled using a binomial distribution with n equal to the total number of troubles (CLEC plus BST troubles.) In this case, the moments for the truncated Z are calculated explicitly using the binomial distribution.

Mean Measures

For mean measures, an adjusted ~~t~~"t" statistic is calculated for each like-to-like cell which has at least 7 BST and 7 CLEC transactions. A permutation test is used when one or both of the BST and CLEC sample sizes is less than 6. Both the adjusted ~~t~~"t" statistic and the permutation calculation are described in ~~the technical~~ Appendix .

Statistical Properties and Definitions.

~~appendix.~~

Ratio Measures

Rules will be given for computing a cell test statistic for a ratio measure, however, the current plan for measures in this category, namely billing accuracy, does not call for the use of a Z parity statistic.

~~APPENDIX D~~

~~Technical Description~~

: Statistical Formulas and Technical Description

We start by assuming that any necessary trimming1 of the data is complete, and that the data are disaggregated so that comparisons are made within appropriate classes or adjustment cells that define "like" observations.

1. Notation and Exact Testing Distributions

Below, we have detailed the basic notation for the construction of the truncated z statistic. In what follows the word "cell" should be taken to mean a like-to-like comparison cell that has both one (or more) ILEC observation and one (or more) CLEC observation.

~~L = the total number of occupied cells~~

L = the total number of occupied cells

~~j = 1,…,L; an index for the cells~~

j = 1,…,L; an index for the cells

~~n_{1j} = the number of ILEC transactions in cell j~~

n1j = the number of ILEC transactions in cell j

~~n_{2j} = the number of CLEC transactions in cell j~~

n2j = the number of CLEC transactions in cell j

~~n_j = the total number transactions in cell j; n_{1j}+ n_{2j}~~

nj= the total number transactions in cell j; n1j+ n2j

~~X_{1jk} = individual ILEC transactions in cell j; k = 1,…, n_{1j}~~

X1jk = individual ILEC transactions in cell j; k = 1,…, n1j

~~X_{2jk} = individual CLEC transactions in cell j; k = 1,…, n_{2j}~~

X2jk = individual CLEC transactions in cell j; k = 1,…, n2j

~~Y_{jk} = individual transaction (both ILEC and CLEC) in cell j~~

$$\cancel{=}\begin{cases} X_{1jk} & k = 1,\ldots,n_{1j} \\ X_{2jk} & k = n_{1j}+1,\ldots,n_j \end{cases}$$

~~$\Phi^{-1}(\cdot)$ = the inverse of the cumulative standard normal distribution function~~

Yjk = individual transaction (both ILEC and CLEC) in cell j

$$=\begin{cases} X & k=1,\ldots,n \\ X & k=n +1,\ldots,n \end{cases}$$

Φ-1( ) = *the inverse of the cumulative standard normal distribution function*

---

1 When it is determined that a measure should be trimmed, a trimming rule that is easy to implement in a production setting is:

For Mean Performance Measures the following additional notation is needed.

~~$\bar{X}_{1j}$ = The ILEC sample mean of cell j~~
of cell j

~~$\bar{X}_{2j}$ = The CLEC sample mean of cell j~~
n of cell

~~s^2_{1j} = The ILEC sample variance in cell j~~
nce in c

~~s^2_{2j} = The CLEC sample variance in cell j~~
nce in c

~~$\{y_{jk}\}$ = a random sample of size n_{2j} from the set of $Y_{j1},\ldots,Y_{jn_j}$; $k = 1,\ldots,n_{2j}$~~

{yjk} = a random sample of size n2j from the set of
$Y ,\ldots, Y$
; k = 1,...,n2j

~~M_j = The total number of distinct pairs of samples of size n_{1j} and n_{2j};~~

$$= \binom{n_j}{n_{1j}}$$

Mj = The total number of distinct pairs of samples of size n1j and n2j;

$$= \binom{n}{n}$$

The exact parity test is the permutation test based on the ~~"modified Z"~~ "modified Z" statistic. For large samples, we can avoid permutation calculations since this statistic will be normal (or Student's t) to a good approximation. For small samples, where we cannot avoid permutation calculations, we have found that the difference between ~~"modified Z"~~ "modified Z" and the textbook ~~"pooled Z"~~ "pooled Z" is negligible. We therefore propose to use the permutation test based on pooled Z for small samples. This decision speeds up the permutation computations considerably, because for each permutation we need only compute the sum of the CLEC sample values, and not the pooled statistic itself.

A permutation probability mass function distribution for cell j, based on the "pooled Z" can be written as

$$\cancel{PM(t) = P(\sum_k y_{jk} = t) = \frac{\textit{the number of samples that sun}}{M_j}}$$

$$CPM(t) = P(\sum y_k \leq t) = \frac{\textit{the number of samples with sun}}{M}$$

and the corresponding cumulative permutation distribution is

$$\cancel{CPM(t) = P(\sum_k y_{jk} \leq t) = \frac{\textit{the number of samples with sum } \leq t}{M_j}}$$

For Proportion Performance Measures the following notation is ~~defined~~
defined

~~a_{1j} = The number of ILEC cases possessing an attribute of interest in cell j~~
a1j = The number of ILEC cases possessing an attribute of interest in cell j

~~a_{2j} = The number of CLEC cases possessing an attribute of interest in cell j~~
a2j = The number of CLEC cases possessing an attribute of interest in cell j

~~a_j = The number of cases possessing an attribute of interest in cell j; $a_{1j}+a_{2j}$~~
aj = The number of cases possessing an attribute of interest in cell j; a1j+ a2j

The exact distribution for a parity test is the hypergeometric distribution. The hypergeometric probability mass function distribution for cell j is

$$\cancel{HG(h) = P(H = h) =} \begin{cases} \dfrac{\binom{n_{1j}}{h}\binom{n_{2j}}{a_j - h}}{\binom{n_j}{a_j}}, & \max(0, a_j - n_{2j}) \le h \le \min(a_j \\ 0 & \text{otherwise} \end{cases}$$

and the cumulative hypergeometric distribution is

$$\cancel{CHG(x) = P(H \le x) =} \begin{cases} 0 & x < \max(0, a_j - n_{2j}) \\ \displaystyle\sum_{h=\max(0,a_j-n_{1j})}^{x} HG(h), & \cancel{\max(0, a_j - n_{2j}) \le x \le \min(a_j, n_{1j}).} \\ 1 & x > \min(a_j, n_{1j}) \end{cases}$$

$$CHG(x) = P(H \le x) = \begin{cases} 0 & x < \max(0, a \; - n \;) \\ \displaystyle\sum HG(h), & \max(0, a \; - n \;) \le x \le \min(a \\ 1 & x > \min(a \;, n \;) \end{cases}$$

For Rate Measures, the notation needed is defined as

~~b_{1j} = The number of ILEC base elements in cell j~~

b1j = The number of ILEC base elements in cell j

~~b_{2j} = The number of CLEC base elements in cell j~~

b2j = The number of CLEC base elements in cell j

~~b_j = The total number of base elements in cell j; $b_{1j}+b_{2j}$~~

bj = The total number of base elements in cell j; b1j+ b2j

~~$\hat{r}_{1j}$ = The ILEC sample rate of cell j; n_{1j}/b_{1j}~~

of c

~~$\hat{r}_{2j}$ = The CLEC sample rate of cell j; n_{2j}/b_{2j}~~

of c

~~q_j = The relative proportion of ILEC elements for cell j; b_{1j}/b_j~~

qj = The relative proportion of ILEC elements for cell j; b1j/bj

The exact distribution for a parity test is the binomial distribution. The binomial probability mass function distribution for cell j is

$$\cancel{BN(k) = P(B = k) =} \begin{cases} \binom{n_j}{k} q_j^k (1-q_j)^{n_j-k}, & 0 \le k \\ 0 & \text{other} \end{cases}$$

and the cumulative binomial distribution is

$$\cancel{CBN(x) = P(B \le x) =} \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{x} BN(k), & 0 \le x \le n_j \\ 1 & x > n_j \end{cases}$$

$$CBN(x) = P(B \le x) = \begin{cases} 0 & x < \\ \sum BN(k), & 0 \le \\ 1 & x > \end{cases}$$

For Ratio Performance Measures the following additional notation is needed.

~~U_{1jk} = additional quantity of interest of an individual ILEC transaction in cell j; k = $1, \ldots, n_{1j}$~~

U1jk = additional quantity of interest of an individual ILEC transaction in cell j; k = 1,…, n1j

~~U_{2jk} = additional quantity of interest of an individual CLEC transaction in cell j; k = $1, \ldots, n_{2j}$~~

U2jk = additional quantity of interest of an individual CLEC transaction in cell j; k = 1,…, n2j

~~$\hat{R}_{ij}$ = the ILEC (I = 1) or CLEC (i = 2) ratio of the total additional quantity of interest to the base transaction total in cell j, i.e., $\sum_k U_{ijk} / \sum_k X_{ijk}$~~

o of

1. Calculating the Truncated Z

The general methodology for calculating an aggregate level test statistic is outlined below.

~~1.~~ Calculate ~~cell weights, W_j.~~ Cell Weights (W_j)

A weight based on the number of transactions is used so that a cell, which has a larger number of transactions, has a larger weight. The actual weight formulae will depend on the type of measure.

Mean or Ratio Measure

$$\cancel{W_j = \sqrt{\frac{n_{1j} n_{2j}}{n_j}}}$$

$$W = \sqrt{}$$

Proportion Measure

$$\cancel{W_j = \sqrt{\frac{n_{2j} n_{1j}}{n_j} \cdot \frac{a_j}{n_j} \cdot \left(1 - \frac{a_j}{n_j}\right)}}$$

$$W = \sqrt{\frac{n\,.\,n}{n}\cdot\frac{a}{n}\cdot\left(1\cdot\right.}$$

$$W = \sqrt{\frac{b\;b}{b}}$$

Rate Measure

$$\cancel{W_j =} \sqrt{\frac{b_{1j}b_{2j}}{b_j}\cdot\frac{n_j}{b_j}}$$

Calculate a Z Value (Z_j) for each Cell

~~2. **In each cell, calculate a Z value, Z_j.**~~ A Z statistic with mean 0 and variance 1 is needed for each cell.

- If $W_j = 0$, set $Z_j = 0$.
- Otherwise, the actual Z statistic calculation depends on the type of performance measure.

Mean Measure

$Z_j = \Phi^{-1}(\alpha)$

where α is determined by the following algorithm.

If $\min(n_{1j}, n_{2j}) > 6$, then determine ~~θ~~α as
~~$\alpha = P(t_{n_{1j}-1}$~~

that is, α is the probability that a t random variable with n_{1j} - 1 degrees of freedom, is less than

$$\cancel{T_j =} \begin{cases} t_j + \frac{g}{6}\left(\frac{n_{1j}+2n_{2j}}{\sqrt{n_{1j}\,n_{2j}(n_{1j}+n_{2j})}}\right)\left(t_j^2 + \frac{n_{2j}-n_{1j}}{n_{1j}+2n_{2j}}\right) & t_j \geq t_{\min j} \\ t_j + \frac{g}{6}\left(\frac{n_{1j}+2n_{2j}}{\sqrt{n_{1j}\,n_{2j}(n_{1j}+n_{2j})}}\right)\left(t_{\min j}^2 + \frac{n_{2j}-n_{1j}}{n_{1j}+2n_{2j}}\right) & \text{otherwise} \end{cases}$$

~~where~~

$$\cancel{t_j =} \frac{\bar{X}_{1j} - \bar{X}_{2j}}{s_{1j}\sqrt{\frac{1}{n_{1j}} + \frac{1}{n_{2j}}}};$$

$$\cancel{t_{\min j} =} \frac{-3\sqrt{n_{1j}n_{2j}n_j}}{g(n_{1j}+2n_{2j})}$$

$$T = \begin{cases} t + \frac{g}{6}\left(\frac{n \;\; + 2n}{\sqrt{n \;\; n \; (n \;\; + n \;)}}\right)\left(t \; + \frac{n \; - n}{n \; + 2n}\right) & t \geq \\ t + \frac{g}{6}\left(\frac{n \;\; + 2n}{\sqrt{n \;\; n \; (n \;\; + n \;)}}\right)\left(t \;\; + \frac{n \; - n}{n \; + 2n}\right) & \text{othe} \end{cases}$$

$$t = \frac{\overline{X} \; -}{s \sqrt{\;}}$$

$$t \;\; = \frac{-3\sqrt{n}}{g(n \;\; +}$$

<u>where</u>

and g is the median value of all values of

$$\gamma_{1j} = \frac{n_{1j}}{(n_{1j}-1)(n_{1j}-2)} \sum_k \left(\frac{X_{1jk} - \overline{X}_{1j}}{s_{1j}} \right)^3$$

$$\gamma = \frac{n}{(n-1)(n-2)} \sum \left(\frac{X - }{s} \right.$$

with

$n > n$

~~with $n_{1j} > n_{3q}$~~ for all values of j. n_{3q} is the 3 quartile of all values of n_{1j}.

Note, that t_j is the "modified Z" statistic. The statistic T_j is a "modified Z" corrected for the skewness of the ILEC data.

If $\min(n_{1j}, n_{2j}) \leq$ 6, and

- ~~a)~~ $M_j \leq$ 1,000 (the total number of distinct pairs of samples of size n_{1j} and n_{2j} is 1,000 or less).

  - Calculate the sample sum for all possible samples of size n_{2j}.
  - Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
  - Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\cancel{\alpha = 1 -} \frac{R_0}{N}$$

- b) $M_j > 1{,}000$

  - Draw a random sample of 1,000 sample sums from the permutation distribution.
  - Add the observed sample sum to the list. There are a total of 1001 sample sums. Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
  - Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\cancel{\alpha = 1 -} \frac{R_0}{1C}$$

$$Z = \frac{n\ a\ - n\ a}{\sqrt{\frac{n\ n\ a\ (n}{n\ -1}}}$$

Proportion Measure

$$Z_j = \frac{n_j\, a_{1j} - n_{1j}\, a_j}{\sqrt{\frac{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)}{n_j - 1}}}$$

Rate Measure

$$\cancel{Z_j} = \frac{n_{1j} - n_j}{\sqrt{n_j\, q_j(1-}}$$

Ratio Measure

$$Z_j = \frac{\hat{R}_{1j} - \hat{R}_{2j}}{\sqrt{V(\hat{R}_{1j})\left(\frac{1}{n_{1j}} + \frac{1}{n_{2j}}\right)}}$$

$$V(\hat{R}_{1j}) = \frac{\sum_k \left(U_{1jk} - \hat{R}_{1j}X_{1jk}\right)^2}{\bar{X}^2_{1j}(n_{1j} - 1)} = \frac{\sum_k U^2_{1jk} - 2\hat{R}_{1j}\sum_k \left(U_{1jk}X_{1jk}\right) + \hat{R}^2_{1j}\sum_k X^2_{1jk}}{\bar{X}^2_{1j}(n_{1j} - 1)}$$

$$Z = \frac{\hat{R} - \hat{R}}{\sqrt{V(\hat{R})\left(\frac{1}{n} + \frac{1}{n}\right)}}$$

$$V(\hat{R}) = \frac{\sum (U - \hat{R}X)}{\bar{X}(n - 1)} = \frac{\sum U - 2\hat{R}\sum(U\,X) + \hat{R}\sum}{\bar{X}(n - 1)}$$

Obtain a Truncated Z Value for each Cell (Z^*j)

~~**3. Obtain a truncated Z value for each cell,**~~ ~~Z^*_j.~~ To limit the amount of cancellation that takes place between cell results during aggregation, cells whose results suggest possible favoritism are left alone. Otherwise the cell statistic is set to zero. This means that positive equivalent Z values are set to 0, and negative values are left alone. Mathematically, this is written as

$$\cancel{Z^*_j = \min(0, Z_j).}$$

$$Z^* = \min($$

Calculate the Theoretical Mean and Variance

~~2.~~Calculate the theoretical mean and variance of the truncated statistic under the null hypothesis of parity,

and

~~$E(Z_j^*|H_0)$ and $Var(Z_j^*|H_0)$. In order to.~~ To compensate for the truncation in step 3, an aggregated, weighted sum of the ~~Z_j^*~~ Z*j will need to be centered and scaled properly so that the final aggregate statistic follows a standard normal distribution.

- If $W_j = 0$, then no evidence of favoritism is contained in the cell. The formulae for calculating

calculating ~~$E(Z_j^*|H_0)$ and $Var(Z_j^*|H_0)$~~ cannot be used. Set both equal to 0.

- If $\min(n_{1j}, n_{2j}) > 6$ for a mean measure,

for a proportion measure,

for a rate measure, or n_{1j} and n_{2j} are large for a ratio measure then

$$\cancel{E(Z_j^*|H_0) = \frac{1}{\sqrt{2\pi}}}, \text{~~and~~}$$

$$\cancel{Var(Z_j^*|H_0) = \frac{1}{2} - \frac{1}{2\pi}};$$

$E(Z|H) = -$

and

$$Var(Z|H) = \frac{1}{2}$$

- Otherwise, determine the total number of values for ~~Z_j^*~~ Z*j. Let z_{ji} and θ_{ji}, denote the values of ~~Z_j^*~~ Z*j and the probabilities of observing each value, respectively.

$$\cancel{E(Z_j^*|H_0) = \sum_i \theta_{ji} z_{ji}}, \text{~~and~~}$$

~~$Var(Z_j^*|H_0) =$~~